足式机器人运动的模型预测路径积分控制采样策略设计

1st ChuyuanTao山东大学人工智能学院中国济南 202599000126 @sdu.edu.cn

2nd FanxinWang*西交利物浦大学机电与机器人系中国苏州 Fanxin.Wang@xjtlu.edu.cn *通讯作者

3rd HaolongJiang西交利物浦大学机电与机器人系中国苏州利物浦大学英国H aolong.Jiang22@student.xjtlu.edu.cn

4th JiaHe西交利物浦大学机电与机器人系中国苏州利物浦大学英国 Jia.He20@student.xjtlu.edu.cn

5th Yiyang Chen 机电与机器人系统 西交利物浦大学 苏州 ，中国 Yiyang.Chen2302@student.xjtlu.edu.c

6nd QingleiBu西交利物浦大学机电与机器人系 西交利物浦大学苏州，中国 n Qinglei.Bu02@xjtlu.edu.cn

arXiv:2601.01409v1 [eess.SY] 4 Jan 2026

摘要——**模型预测路径积分（MPPI）控制已成为一种强大的基于采样的最优控制方法，适用于复杂、非线性和高维系统。然而，将 MPPI 直接应用于足式机器人系统面临着若干挑战。本文系统地研究了 MPPI 框架中采样策略设计对足式机器人运动的作用。基于结构化控制参数化的思路，我们在该框架内探索并比较了多种采样策略，包括非结构化方法和基于样条的方法。通过在四足机器人平台上开展大量仿真，我们评估了不同采样策略对控制平滑性、任务性能、鲁棒性和采样效率的影响。研究结果为在复杂足式系统中部署 MPPI 时采样设计的实际影响提供了新的见解。**

索引词——**模型预测路径积分控制，基于采样的最优控制，足式机器人运动，结构化控制参数化，四足机器人仿真**

I. 引言

近年来，MPPI 控制已成为一种高效的基于采样的最优控制框架，可用于处理复杂的非线性高维系统 [1]–[3]。MPPI 利用蒙特卡罗滚动采样和重要性采样，无需显式计算梯度，并且能够自然地适应非线性动力学、非凸代价函数以及难以建模的交互作用。MPPI 的一大优势是其天然适用于并行计算架构（例如图形处理器），因此可以同时评估数千个轨迹滚动采样 [4], [5]。这一特性使 MPPI 尤其适用于具有复杂动力学和混合行为的系统，包括腿式机器人；在这些系统中，解析解和凸 formulation 往往并不实用 [6], [7]。

尽管具有这些优势，将 MPPI 直接应用于足式机器人系统仍面临若干挑战。一个主要限制在于控制采样过程本身。标准的 MPPI 实现通常在每个时间步独立采样控制输入，这可能导致控制序列在时间上不相关且不平滑 [8]。对于在接触丰富环境中运行的腿式机器人，这种不平滑的控制信号可能导致行为不稳定、执行器 effort 过大或接触一致性较差。足式运动动力学具有混合特性，这进一步加剧了上述问题：接触切换和碰撞事件引起的不连续性给控制与优化带来了重大挑战 [9], [10]。

为应对这些困难，先前的研究在基于采样的控制和 MPC框架中探索了结构化控制参数化表示。研究表明，基于样条、基函数或降阶参数化的轨迹表示能够提升最优控制中的平滑性和采样效率 [11], [12]。值得注意的是，论文 “采用模型预测路径积分控制的腿式机器人实时全身控制” [13] 通过使用三次样条参数化，在较低维空间中采样控制轨迹，实现了腿式机器人上的成功实时MPPI实现。作者没有在每个时间步直接采样离散控制动作，而是使用样条系数对控制序列进行参数化，从而保证时间平滑性并降低控制

采样空间的有效维度。该方法显著提升了控制平滑性，并使机器人能够在真实世界实验中实现稳定的全身行为。

受这一研究方向的启发，本文系统地探讨了采样策略设计在足式机器人控制中 MPPI 的作用。基于结构化控制参数化的思想，我们在 MPPI 框架内探索并比较了多种采样策略，包括非结构化方法和基于样条的方法。通过在足式机器人平台上进行大量仿真，我们评估了不同采样策略对控制平滑性、任务性能、鲁棒性和采样效率的影响。结果揭示了采样设计对基于 MPPI 的足式运动的实际影响，并突显了采样策略是将 MPPI 部署于复杂机器人系统时的关键因素。

A. 相关工作

针对在非线性动力学、不确定性和安全约束下运行的机器人系统，基于采样的最优控制方法已得到广泛研究 [3], [14]。在这些方法中，基于路径积分的控制方法（例如模型预测路径积分（MPPI））凭借通过随机采样和并行计算处理非凸最优控制问题的能力，受到广泛关注 [1], [15]。近期研究着重于将控制理论结构融入采样过程，以提升路径积分控制的安全性和鲁棒性。Tao 等人 [16] 提出了一种将随机控制屏障函数（控制障碍函数）纳入路径积分控制的框架，在保持计算效率的同时实现概率安全保证。同样，Wang 等人 [17]提出了 DBaS-Log-MPPI，通过引入屏障状态来塑造采样分布，从而提高受约束环境中的安全性和效率。

与此同时，信息引导采样也被证明对于复杂环境中的运动规划和基于学习的控制至关重要。Tao等人 [18] 提出了一种用于动态且障碍物密集环境的引导式基于采样的运动规划算法，凸显了采样过程中的结构如何显著提升规划效率和鲁棒性。强化学习领域也出现了相关思路，通过自适应控制和基于障碍函数的机制来保障安全性和鲁棒性 [19],[20]。尽管基于学习的方法通常依赖离线训练，但它们与基于采样的最优控制有一个共同见解：对动作空间或控制空间施加结构对于实现可靠性能至关重要。尤其对于足式机器人，先前研究表明，直接对时间离散化控制输入进行采样往往会导致不平滑且不稳定的行为，因此在模型预测控制和 MPPI 实现中，促使人们采用样条或降阶表示等结构化参数化方式 [7], [13]。不同于引入单一结构化表示，本文在 MPPI 框架中系统比较了多种采样策略，以更好地理解它们对足式运动任务中性能和鲁棒性的影响。

II. 预备知识

A. 实验平台与任务

本文的所有实验均在一台宇树四足机器人上进行，该机器人配备力矩控制关节和机载状态估计功能。机器人可提供完整状态反馈，包括基座位姿、速度和关节状态，从而能够在动态运动任务中进行实时全身控制。

我们在模型预测路径积分（MPPI）框架下，针对一组具有代表性的运动场景评估采样策略。这些场景着重考察控制平滑性、适应性和鲁棒性的不同方面：

- **平地行走**：在平坦地形上进行稳态运动，用于评估标称跟踪性能和控制平滑性。
- **爬楼梯**：穿越具有离散高度变化的楼梯，需要精准的足部落点和协调的全身运动。
- **大型障碍物（箱体）跨越**：跨越大型箱体障碍物，这会导致接触构型发生突变，并要求控制轨迹具有较强的表达能力。

这些任务常用作腿式机器人运动的基准测试，并共同涵盖稳定、中等动态以及高度非平滑的交互模式。所有采样策略均在相同的任务设置、代价函数和控制约束下进行评估，以确保比较公平。

B. 模型预测路径积分控制

模型预测路径积分（MPPI）控制是一种基于采样的随机最优控制方法，通过评估采样轨迹的期望代价来计算控制更新。MPPI 采用滚动时域方式运行，尤其适用于腿式机器人这类高维非线性系统。

令 $x_t \in \mathbb{R}^n$ 表示系统状态，$u_t \in \mathbb{R}^m$ 表示控制输入。给定长度为 H 的时域和一个标称控制序列 $\bar{u}_{0:H-1}$，MPPI 通过向控制序列注入噪声来生成 N 个随机展开：

$$u_t^{(k)} = \bar{u}_t + \epsilon_t^{(k)}, \quad \epsilon_t^{(k)} \sim \mathcal{N}(0, \Sigma), \tag{1}$$

其中 $k = 1, \dots, N$对滚动采样进行索引。

每个轨迹展开都使用系统动力学向前模拟，以获得状态轨迹 $\{x_t^{(k)}\}_{t=0}^{H}$。累积代价计算如下：

$$S^{(k)} = \sum_{t=0}^{H-1} \ell(x_t^{(k)}, u_t^{(k)}) + \phi(x_H^{(k)}), \tag{2}$$

其中 $\ell(\cdot)$ 和 $\phi(\cdot)$ 分别表示运行代价和终端代价。

标称控制序列通过对噪声采样结果进行成本加权平均来更新：

$$\bar{u}_t \leftarrow \bar{u}_t + \sum_{k=1}^{N} w^{(k)} \epsilon_t^{(k)}, \quad w^{(k)} = \frac{\exp(-S^{(k)}/\lambda)}{\sum_{j=1}^{N} \exp(-S^{(j)}/\lambda)}, \tag{3}$$

算法 1模型预测路径积分控制

要求： 初始状态 x_0，初始名义控制 $\bar{u}_{0:H-1}$

对于每次控制迭代**执行**
  对于 $k = 1$ 到 N **执行**
    采样控制扰动 $\epsilon^{(k)}$
    构建控制轨迹 $u^{(k)} = \bar{u} + \epsilon^{(k)}$
    展开系统动力学以获得 $\{x_t^{(k)}\}_{t=0}^{H}$
    计算轨迹代价 $S^{(k)}$
  循环结束
  计算重要性权重 $w^{(k)}$
  更新名义控制 $\bar{u}_{0:H-1}$
  执行第一个控制输入并平移时域
结束循环

其中，$\lambda > 0$ 是控制加权锐度的温度参数。

C. MPPI 在腿式机器人全身控制中的应用

本研究基于文献 [13],中提出的 MPPI 全身控制框架。在该框架中，MPPI 直接应用于腿式机器人的高维关节空间控制。该框架采用三次样条参数化来降低有效采样维度并确保控制轨迹平滑，从而支持在硬件上实时运行。

受这一方法启发，我们采用相同的 MPPI 结构、代价设计和滚动时域执行策略。不过，我们不再将采样限制为三次样条参数化，而是研究多种结构化采样策略，包括三次样条、B´ézier 曲线和线性插值参数化。我们的目标是分离并量化采样结构对 MPPI 在足式运动任务中性能的影响。

重要的是，本文考虑的所有方法都采用相同的 MPPI更新规则，区别仅在于如何将随机扰动映射为连续时间控制轨迹。

算法 1 总结了本文所采用的 MPPI 流程。

III. MPPI 算法中的不同采样策略

A. 三次样条采样

为了生成维度更低的平滑控制轨迹，我们采用三次样条采样策略。我们不在每个时间步对控制输入进行采样，而是在少量节点处采样控制值，然后使用三次样条插值重建完整轨迹。

令 H 表示规划时域，m 表示动作维度。我们选择 K 个节点索引

$$\mathcal{T} = \{t_1, t_2, \ldots, t_K\}, \quad 0 = t_1 < t_2 < \cdots < t_K = H - 1, \tag{4}$$

这些索引在整个时域内均匀分布。

给定名义控制轨迹 $\bar{u}_{0:H-1} \in \mathbb{R}^{H\times m}$，相应的名义节点值为

$$\bar{u}_{t_i} = \bar{u}(t_i), \quad i = 1, \ldots, K. \tag{5}$$

对于每个轨迹展开 k，在节点处施加独立的高斯噪声：

$$\epsilon_{i,j}^{(k)} \sim \mathcal{N}(0, \sigma_j^2), \quad i = 1, \ldots, K,\ j = 1, \ldots, m, \tag{6}$$

其中，σ_j 表示第 j 个动作维度的噪声标准差。

扰动后的节点值为

$$u_{t_i}^{(k)} = \bar{u}_{t_i} + \epsilon_i^{(k)}. \tag{7}$$

对于每个动作维度 j，通过对扰动后的节点值进行插值来构造三次样条：

$$u_j^{(k)}(t) = \mathrm{CubicSpline}\big(\{(t_i, u_{t_i,j}^{(k)})\}_{i=1}^{K}\big), \quad t = 0, \ldots, H-1. \tag{8}$$

将所有动作维度堆叠起来，即可得到采样的控制轨迹

$$u^{(k)}(t) = \big[u_1^{(k)}(t), \ldots, u_m^{(k)}(t)\big]^\top. \tag{9}$$

最后，通过逐元素裁剪来施加执行器限制：

$$u^{(k)}(t) \leftarrow \mathrm{clip}\big(u^{(k)}(t), u_{\min}, u_{\max}\big). \tag{10}$$

三次样条采样可确保控制轨迹的 C^2 连续性，并将采样维度从 $H \times m$ 降至 $K \times m$。

B. 贝´塞尔曲线采样

为生成具有显式低维参数化的平滑控制轨迹，我们采用贝´塞尔曲线采样。与基于样条的采样类似，该方法扰动一组精简的控制点，并通过确定性插值重构完整的控制轨迹。与三次样条不同，贝´塞尔曲线通过伯恩斯坦多项式基提供全局参数化。

令 H 表示规划时域，m 表示动作维度。我们均匀选择 K 个控制点索引

$$\mathcal{T} = \{t_1, t_2, \ldots, t_K\}, \quad 0 = t_1 < t_2 < \cdots < t_K = H - 1, \tag{11}$$

使其分布于整个时域内。

给定名义控制轨迹 $\bar{u}_{0:H-1} \in \mathbb{R}^{H\times m}$，名义控制点定义为

$$\bar{p}_i = \bar{u}(t_i), \quad i = 1, \ldots, K. \tag{12}$$

对于每个轨迹展开 k，独立地向每个控制点施加高斯噪声：

$$\epsilon_{i,j}^{(k)} \sim \mathcal{N}(0, \sigma_j^2), \quad i = 1, \ldots, K,\ j = 1, \ldots, m. \tag{13}$$

扰动后的控制点表示为

$$p_i^{(k)} = \bar{p}_i + \epsilon_i^{(k)}. \tag{14}$$

使用伯恩斯坦多项式基函数重建完整的控制轨迹。令 $n = K - 1$ 表示 B´ézier 曲线的次数，$\tau \in [0, 1]$ 表示归一化时间变量。伯恩斯坦基定义为

$$B_i^n(\tau) = \binom{n}{i} \tau^{\,i} (1-\tau)^{n-i}, \quad i = 0, \ldots, n. \tag{15}$$

轨迹展开 k 的采样控制轨迹为

$$u_j^{(k)}(t) = \sum_{i=0}^{n} p_{i,j}^{(k)} \, B_i^n\Big(\tfrac{t}{H-1}\Big), \quad t = 0, \dots, H-1, \tag{16}$$

对每个动作维度 j。

将所有动作维度堆叠起来，得到

$$u^{(k)}(t) = \big[u_1^{(k)}(t), \dots, u_m^{(k)}(t)\big]^\top. \tag{17}$$

最后，通过裁剪来施加执行器限制：

$$u^{(k)}(t) \leftarrow \mathrm{clip}\big(u^{(k)}(t), u_{\min}, u_{\max}\big). \tag{18}$$

贝塞尔曲线采样将采样维度从$H \times m$ 降低至 $K \times m$，同时确保控制轨迹的全局平滑性。与三次样条采样相比，贝塞尔曲线通过显式控制点提供了更直观的参数化，并保证轨迹位于这些点构成的凸包内。

C. 基于线性插值的采样

除基于样条的采样外，我们还在 MPPI 框架中采用基于线性插值的采样策略，作为一种低复杂度基线。其目标是在保留控制轨迹基本时间结构的同时，降低采样空间的维度，而不施加高阶平滑性约束。

令控制时域为 T，动作维度为 d_u。我们不直接对完整控制序列$\mathbf{u}_{0:T-1} \in \mathbb{R}^{T \times d_u}$进行采样，而是定义一组由以下索引标记的 K航点

$$\mathcal{I} = \{i_1, i_2, \dots, i_K\}, \quad i_k \in \{0, \dots, T-1\}, \tag{19}$$

其中，各索引在整个时域内均匀分布：

$$i_k = \left\lfloor \frac{(k-1)(T-1)}{K-1} \right\rceil, \quad k = 1, \dots, K. \tag{20}$$

从当前均值控制轨迹 $\bar{\mathbf{u}}$中提取均值航点，具体为

$$\bar{\mathbf{w}}_k = \bar{\mathbf{u}}_{i_k}, \quad k = 1, \dots, K. \tag{21}$$

对于每次轨迹展开 $n \in \{1, \dots, N\}$，根据以下方式独立采样航点扰动：

$$\mathbf{w}_k^{(n)} = \bar{\mathbf{w}}_k + \boldsymbol{\epsilon}_k^{(n)}, \quad \boldsymbol{\epsilon}_k^{(n)} \sim \mathcal{N}(\mathbf{0}, \Sigma), \tag{22}$$

其中，$\Sigma = \sigma^2 I$ 表示各向同性协方差。

通过在相邻航点之间进行分段线性插值，重建完整的控制轨迹：

$$\mathbf{u}_t^{(n)} = \mathrm{Interp}\Big(t; \, \{(i_k, \mathbf{w}_k^{(n)})\}_{k=1}^{K}\Big), \quad t = 0, \dots, T-1, \tag{23}$$

插值在每个动作维度上独立进行。通过逐元素裁剪来约束动作范围：

$$\mathbf{u}_t^{(n)} \leftarrow \mathrm{clip}(\mathbf{u}_t^{(n)}, \mathbf{u}_{\min}, \mathbf{u}_{\max}). \tag{24}$$

这种方法将随机采样维度从 $T \times d_u$ 降至 $K \times d_u$。然而，由于缺乏曲率连续性，生成的轨迹不如基于样条的参数化所生成的轨迹平滑。

IV. 仿真结果：采样策略比较

本节对 MPPI 框架内不同的控制采样参数化进行比较评估。所有实验均使用宇树级别的四足机器人模型在仿真中进行，旨在单独考察采样策略对运动性能的影响。

A. 实验设置

上一节所述的 MPPI 控制器在所有实验中均未作任何修改。具体而言，所有方法使用相同的系统动力学模型、代价函数设计、温度参数、规划时域、滚动采样数量和控制频率。实验之间唯一的差异在于控制扰动如何参数化，并在预测时域内如何重构。

每种控制策略均在三个难度递增的代表性运动任务上进行评估：

1) **平地行走**：在平坦地形上朝目标位置直线行走。2)**爬楼梯**：沿具有离散高度变化的楼梯向上攀爬，需要持续的垂直运动和精准的足部落点。3)**大型障碍物跨越**：跨越一个高箱子，该障碍会导致机器人躯干与腿部之间产生显著的运动学和动力学耦合。

对于每项任务和每种采样策略，均在相同的初始条件下进行五次独立试验。如果机器人未摔倒且未违反关节或接触约束，成功到达目标位置，则该试验被视为成功。

我们采用三项性能指标：

- 成功率：成功到达目标的试验所占的百分比。
- 到达目标的步数：到达目标所需的控制步数，并对成功试验取平均值。
- 计算时间：在相同硬件上测得的每次 MPPI 迭代的平均实际运行时间。

这些指标共同反映了鲁棒性、控制效率和计算成本，这些因素对于实时全身控制都至关重要。

B. 平地行走

首先，我们在平地行走任务中评估所提出的采样参数化表示，其中四足机器人被命令沿直线行走 1.0 m。该场景作为基准测试，用于评估在不存在环境扰动的情况下的步态稳定性、收敛速度和计算效率。

表 I 汇总了每种方法五次试验的定量结果。采用独立高斯采样（正态分布）的基线 MPPI 成功率仅为 60%，且收敛速度最慢，需要 1358.0 ± 571.8步才能到达目标。步数的巨大方差反映了

轨迹展开不稳定，关节指令呈振荡状态，因此经常无法实现可靠的向前运动。

相比之下，强制保持时间连贯性的结构化采样方法，能显著提升鲁棒性和收敛速度。其中，*CubicSpline-k4* 的成功率达到 100%，收敛速度也显著快于其他所有方法，仅用 444.0 ± 30.2 步便到达目标。这一结果表明，低维样条参数化能有效平衡表达能力与正则化，从而生成平滑、稳定的控制序列，并使不同试验间的方差降至最低。

然而，提高样条的维度会降低鲁棒性。尽管 *CubicSpline-k8* 在收敛速度方面仍优于基线方法，但其成功率降至 80%，步数方差也显著增加（687.0±661.0），这表明在采样预算有限时，该方法容易受到过度参数化的影响。这一趋势凸显了在 MPPI 中控制动作空间有效维度的重要性。

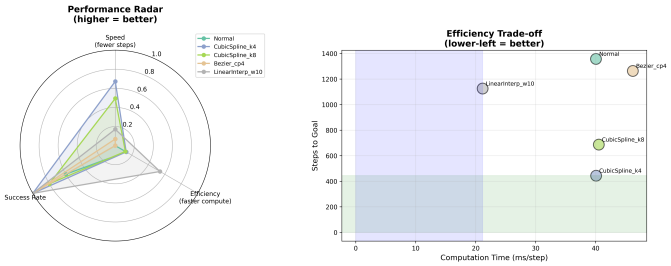


图 1. 平地行走的性能雷达图与效率权衡。

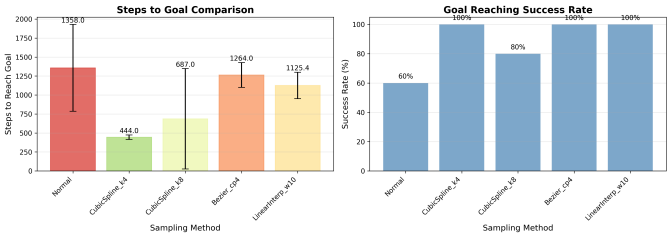


图 2. 不同采样方法在平地行走中的到达目标所需步数与成功率对比。

不同的结构化表示各有取舍。*Bezier-cp4* 和 *LinearInterp-w10*的成功率均为 100%，但平均需要超过 1100 步才能到达目标。值得注意的是，*LinearInterp-w10*是计算效率最高的方法，每步平均计算时间为21.14 ± 4.17 毫秒，约为基于样条方法的一半，但代价是收敛较慢。

图 2 进一步说明了这些权衡关系，该图展示了计算时间与收敛速度之间的关系；图 1 则提供了涵盖速度、效率和成功率的归一化性能雷达图对比。总体而言，结果表明，中等维度的样条参数化表示，尤其是 *CubicSpline-k4*，在平地行走任务中，能够在收敛速度、鲁棒性和计算成本之间实现最有利的平衡。

表 I 平地行走性能比较。

方法	成功率（%）	到达目标的步数	时间（毫秒）
正态分布	60	1358.0 ± 571.8	40.05 ± 13.23
CubicSpline-k4	100	444.0 ± 30.2	40.07 ± 8.70
CubicSpline-k8	80	687.0 ± 661.0	40.52 ± 11.21
Bezier-cp4	100	1264.0 ± 163.8	46.19 ± 13.54
LinearInterp-w10	100	1125.4 ± 174.7	21.14 ± 4.17

C. 爬楼梯

接下来，我们在爬楼梯任务上评估这些采样策略：四足机器人需要在 3.3 m 的前进距离内登上楼梯。与平地行走相比，该场景对长时域协调、精确的落足点布置以及身体垂直运动提出了显著更高的要求，因此成为检验采样动作序列时间一致性的严格测试。

表 II 报告了每种方法进行五次试验的定量性能。采用独立高斯采样（正态分布）的基线 MPPI 完全失败，成功率为 0%，并且始终达到 4000 步的最大回合长度。尽管其每步计算时间相对较短，但由于缺乏时间结构，生成的动作序列不连贯，无法维持稳定的爬楼梯行为。*Bezier-cp4* 也表现出类似的失败模式，其成功率同样为 0 %，同时在所有方法中产生了最高的计算成本。

相比之下，基于样条的采样在这项长时域任务中展现出显著优势。*CubicSpline-k4* 和*CubicSpline-k8* 均达到 100% 的成功率，平均约用 1000 步即可可靠地完成楼梯攀爬。其中，*CubicSpline-k4* 收敛稍快，用时 1010.6 ± 232.8 步，而*CubicSpline-k8*为 1031.6 ± 332.4 步，同时方差也更低。这一结果再次表明，中等维度的样条参数化足以灵活表达复杂动作，同时不会对采样噪声产生不必要的敏感性。

LinearInterp-w10 方法取得了一定成功，成功率为 60%，但需要明显更多步数才能到达目标（2817.2 ± 1032.3），且不同试验之间的结果波动较大。虽然其较低的计算时间使其在效率方面颇具吸引力，但这种分段线性结构似乎不足以捕捉爬楼梯所需的平滑、协调运动。

图 4 展示了计算时间与收敛速度之间的权衡，清晰地区分了成功的基于样条的方法与失败案例。图 3 进一步总结了收敛速度、效率和成功率方面的归一化性能。总体而言，这些结果表明，具有时间平滑性且维度较低的样条参数化表示

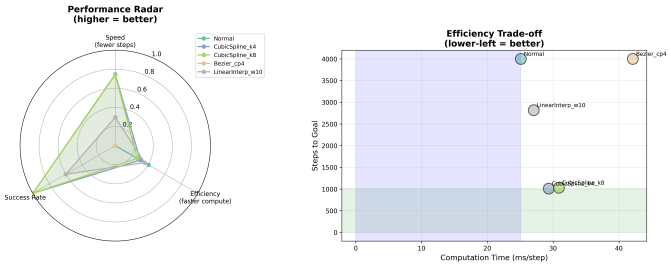


图 3. 爬楼梯任务的性能雷达图与效率权衡。

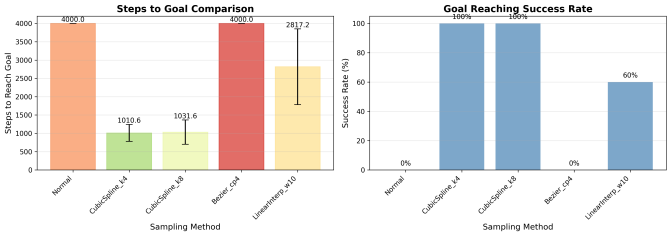


图 4. 爬楼梯任务中不同采样方法的到达目标所需步数和成功率比较。

不仅有益，而且对于在 MPPI 框架内解决诸如爬楼梯之类的接触密集型长时域 locomotion 任务至关重要。

D. 大型箱体障碍物跨越

最后，我们在大型障碍物跨越任务中评估采样策略：四足机器人必须跨过一个高箱体，同时向前行进 1.0 米。与爬楼梯不同，该场景强调短时域、高度耦合的全身协调，要求机器人在有限的空间范围内快速调整身体姿态和足部落点。

表 III 总结了每种方法进行五次试验的性能。采用独立高斯采样（正态分布）的基线 MPPI 在所有试验中再次失败，成功率为 0%，且每次都达到 4000 步的最大回合长度。尽管计算时间适中，但缺乏结构化的时间耦合，使控制器无法产生跨越障碍所需的协调抬腿和落地动作。

在所有方法中，*CubicSpline-k4*的整体性能最佳，成功率达到 100%，且到达目标所需步数最少（955.4 ± 739.8）。尽管其方差高于平地行走时的方差，但这是障碍物跨越过程中接触事件本质上不连续所致，符合预期。值得注意的是，*CubicSpline-k4* 兼具鲁棒性和计算效率，平均每步计算时间为23.25 ± 5.01 毫秒。

增加样条维度会导致该任务的鲁棒性降低。*CubicSpline-k8*的成功率为 80%，且平均需要更多步数（1332.4 ± 1379.3），尽管就每步计算时间而言，它是最快的方法。这一结果表明，虽然高维参数化表示可能带来额外的灵活性，但也可能

爬楼梯性能对比，表 II。

方法	成功率（%）	到达目标的步数	时间（ms）
正态分布	0	4000.0 ± 0.0	25.08 ± 5.79
CubicSpline-k4	100	1010.6 ± 232.8	29.34 ± 6.89
CubicSpline-k8	100	1031.6 ± 332.4	30.87 ± 11.61
Bezier-cp4	0	4000.0 ± 0.0	42.11 ± 11.94
LinearInterp-w10	60	2817.2 ± 1032.3	27.04 ± 8.52

在短时域、接触丰富的机动过程中引入不必要的变异性。

其余结构化方法的表现处于中等水平。*LinearInterp-w10* 成功率为 60%，但步数方差较大（2239.4 ± 1461.0），表明其应对障碍物时表现不稳定。*Bezier-cp4* 在结构化方法中表现最差，仅在 40% 的试验中成功，而且常常无法生成足够激进或时机恰当的动作来跨越障碍物。

图6展示了收敛速度与计算成本之间的权衡，清楚地区分了成功的基于样条的方法与失败案例。图5进一步对速度、效率和成功率进行了归一化比较。综合来看，这些结果表明，对于时域较短但需要高度协调的行为（例如跨越大型障碍物），在 MPPI 中，低维样条参数化表示再次展现出表达能力、鲁棒性与采样效率之间的最佳平衡。

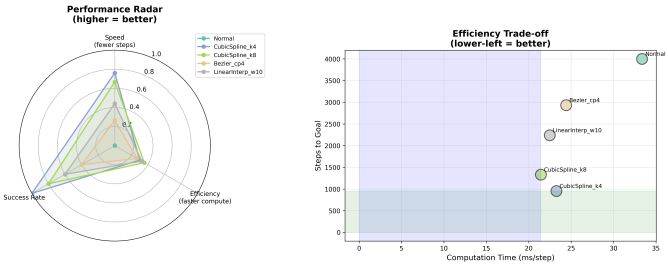


图5. 大型箱体障碍物跨越任务的性能雷达图与效率权衡。

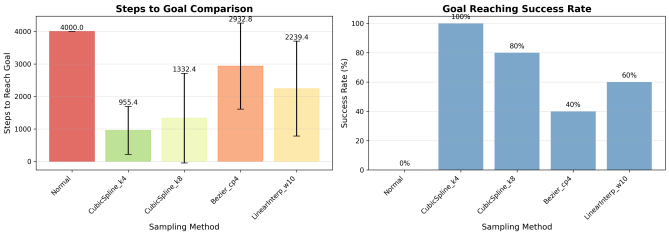


图6. 不同采样方法在爬楼梯任务中的到达目标所需步数和成功率对比。

E. 仿真总结

总而言之，仿真结果表明，在 MPPI 框架内选择何种采样参数化表示，会显著影响任务性能

表 III 大箱体穿越性能比较。

方法	成功率 (%)	到达目标的步数	时间 (ms)
正态分布	0	4000.0 ± 0.0	33.36 ± 8.79
CubicSpline-k4	100	955.4 ± 739.8	23.25 ± 5.01
CubicSpline-k8	80	1332.4 ± 1379.3	21.43 ± 3.89
Bezier-cp4	40	2932.8 ± 1323.5	24.40 ± 4.39
LinearInterp-w10	60	2239.4 ± 1461.0	22.47 ± 4.58

以及在不同运动场景中的鲁棒性。尽管所有受评估的策略都能在平坦地形上生成可行运动，但在爬楼梯和大型障碍物跨越等更具挑战性的环境中，策略之间的差异便十分明显。能够更好地捕捉控制序列内在结构的参数化表示，始终可以实现更低的任务代价、更高的跟踪精度和更小的控制变化性，尤其是在接触丰富且高度非线性的动力学条件下。此外，这些策略具有更高的采样效率，无需增加计算预算即可实现更稳定的行为。综合来看，这些发现表明，信息引导采样设计是将 MPPI 扩展到复杂足式运动任务的关键因素，并促使我们在四足平台上部署 MPPI 时采用结构化或低维采样空间。

V. 结论与未来工作

本文研究了 MPPI 框架中采样策略设计在足式机器人运动中的作用。我们在四足机器人上进行了大量仿真，测试并比较了采用不同控制采样策略的 MPPI控制器在多种地形和任务场景中的表现。结果表明，尽管 MPPI 提供了统一的随机最优控制框架，但其在腿式机器人上的实际性能会受到采样参数化选择的显著影响。具体而言，与非结构化采样方法相比，结构化采样策略能够提升任务表现、增强在接触丰富环境中的鲁棒性，并提高采样效率。这些发现凸显了采样设计是有效将 MPPI 应用于复杂足式运动问题的关键因素。

基于本研究所得的见解，未来工作将着眼于更广泛地研究最优控制中的采样策略。我们旨在分析不同的采样参数化表示如何影响随机最优控制算法（包括但不限于 MPPI）的探索–利用权衡、收敛性质和计算效率。此外，我们计划探索自适应和基于学习的采样机制，使其能够根据任务需求和系统动力学在线调整。最终，这一研究方向旨在为采样设计建立系统化的指导原则，从而支持基于采样的最优控制方法在真实世界机器人系统中的可扩展、可靠部署。

参考文献

[1] G. Williams, N. Wagener, B. Goldfain, P. Drews, J. M. Rehg, B. Boots, 和 E. A. Theodorou，“面向基于模型的强化学习的信息论模型预测控制”，载于《2017年IEEE国际机器人与自动化会议（ICRA）》，第1714–1721页，IEEE，2017年。[2] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, 和 E. A. Theodorou，“信息论模型预测控制：理论及其在自动驾驶中的应用”，《IEEE机器人学汇刊》，第34卷，第6期，第1603–1622页，2018年。[3] E. Theodorou, J. Buchli, 和 S. Schaal，“一种用于强化学习的广义路径积分控制方法”，《机器学习研究杂志》，第11卷，第3137–3181页，2010年。[4] G. Williams, A. Aldrich, 和 E. Theodorou，“使用协方差变量重要性采样的模型预测路径积分控制”，arXiv预印本 arXiv:1509.01149，2015年。[5] K. Lowrey, A. Rajeswaran, S. Kakade, E. Todorov, 和 I. Mordatch，“在线规划，离线学习：通过基于模型的控制实现高效学习与探索”，arXiv预印本 arXiv:1811.01848，2018年。[6] M. Hutter, C. Gehring, D. Jud, A. Lauber, C. D. Bellicoso, V. Tsounis, J. Hwangbo, K. Bodie, P. Fankhauser, M. Bloesch 等，“ANYmal——一款高度灵活且动态性能出色的四足机器人”，载于《2016年IEEE/RSJ智能机器人与系统国际会议（IROS）》，第38–44页，IEEE，2016年。[7] N. Rathod, A. Bratta, M. Focchi, M. Zanon, O. Villarreal, C. Semini, 和 A. Bemporad，“面向足式运动的地形感知模型预测控制”，2021年。[8] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, 和 E. A. Theodorou，“使用模型预测路径积分控制实现激进驾驶”，载于《2016年IEEE国际机器人与自动化会议（ICRA）》，第1433–1440页，IEEE，2016年。[9] I. Chatzinikolaidis, Y. You, 和 Z. Li，“利用解析可解接触模型进行接触隐式轨迹优化，以实现可变地面上的运动”，《IEEE机器人与自动化快报》，第5卷，第4期，第6357–6364页，2020年。[10] S. Gangapurwala, M. Geisert, R. Orsolino, M. Fallon, 和 I. Havoutis，“RLOC：结合强化学习与最优控制的地形感知足式运动”，《IEEE机器人学汇刊》，第38卷，第5期，第2908–2927页，2022年。[11] J. T. Betts，《使用非线性规划进行最优控制与状态估计的实用方法》。SIAM，2010年。[12] Y. Tassa, N. Mansard, 和 E. Todorov，“控制受限微分动态规划”，载于《2014年IEEE国际机器人与自动化会议（ICRA）》，第1168–1175页，IEEE，2014年。[13] J. Alvarez-Padilla, J. Z. Zhang, S. Kwok, J. M. Dolan, 和 Z. Manchester，“使用模型预测路径积分控制对腿式机器人进行实时全身控制”，载于《2025年IEEE国际机器人与自动化会议（ICRA）》，第14721–14727页，IEEE，2025年。[14] E. Todorov，“最优动作的高效计算”，《美国国家科学院院刊》，第106卷，11483，2009年28，11478，第–页。[15] M. Testouri, G. Elghazaly, 和 R. Frank，“面向自动驾驶系统的安全实时运动规划框架：一种MPPI方法”，arXiv预印本 arXiv:2308.01654，2023年。[16] C. Tao, H.-J. Yoon, H. Kim, N. Hovakimyan, 和 P. Voulgaris，“结合随机控制屏障函数的路径积分方法”，载于《2022年IEEE第61届决策与控制会议（CDC）》，第1654–1659页，IEEE，2022年。[17] F. Wang, H. Jiang, C. Tao, W. Wan, 和 Y. Cheng，“DBAS-LOG-MPPI：通过屏障状态实现高效且安全的轨迹优化”，arXiv预印本 arXiv:2504.06437，2025年。[18] C. Tao, H. Kim, H.-J. Yoon, P. Voulgaris, 和 N. Hovakimyan，“面向动态及障碍物密集环境的引导式基于采样的运动规划算法”，《制导、控制与动力学杂志》，第1–18页，2025年。[19] Y. Cheng, P. Zhao, F. Wang, D. J. Block, 和 N. Hovakimyan，“提高强化学习策略的鲁棒性：使用 $\mathcal{L}_1$ 自适应控制”，《IEEE机器人与自动化快报》，第7卷，第3期，第6574–6581页，2022年。[20] Y. Cheng, P. Zhao, 和 N. Hovakimyan，“利用基于扰动观测器的控制屏障函数实现安全高效的强化学习”，载于《学习动力学与控制会议》，第104–115页，PMLR，2023年。